

Exploratory Data Analysis & Vendor Summary Report

This report outlines the transformation of raw inventory and sales data into a consolidated **Vendor Sales Summary**. The primary objective was to facilitate data-driven decisions regarding **vendor profitability** and **product pricing optimization**.

1. Data Source Overview

The initial analysis involved seven tables from the inventory.db database. The dataset is substantial, with the sales table containing over **12.8 million records** and the purchases table exceeding **2.3 million records**.

Table Name	Key Contents	Purpose
begin_inventory / end_inventory	Hand-on quantity & prices	Tracking stock levels over time.
purchases	Vendor, PO, Date, Purchase Price	Tracking what was bought and at what cost.
purchase_prices	Brand, Standard Price, Volume	Reference for actual product market prices.
sales	Sales Quantity, Dollars, Excise Tax	Tracking revenue and consumer demand.
vendor_invoice	Freight, PO Number, Totals	Accounts payable and logistics costs.

2. The Consolidation Strategy

To enable profitability analysis, we engineered a master view: vendor_sales_summary. This required complex SQL Common Table Expressions (CTEs) to join purchasing data with actual sales performance and freight overhead.

Key Metrics Calculated:

- **Gross Profit:** Total Sales Dollars minus Total Purchase Dollars.
- **Profit Margin:** Percentage of sales revenue exceeding purchase costs.
- **Stock Turnover:** Total Sales Quantity divided by Total Purchase Quantity (measures inventory efficiency).
- **Sales-to-Purchase Ratio:** Revenue generated per dollar spent on stock.

3. Key Findings from the Summary

The pre-aggregated table reveals significant insights into high-performing vendors:

- **Top Revenue Drivers:** Vendors like **BROWN-FORMAN CORP** and **MARTIGNETTI COMPANIES** dominate in total purchase and sales volume, particularly with flagship products like *Jack Daniels No 7* and *Tito's Handmade Vodka*.
- **Profitability Variance:** While some brands show strong margins, the analysis identified a minimum Gross Profit of **-\$52,002.78**, indicating specific areas where purchase costs significantly outweigh sales revenue.
- **Efficiency Metrics:**
 - **Stock Turnover:** High-volume items (e.g., *Absolut 80 Proof*) show turnover ratios near **1.0**, suggesting a healthy balance between buying and selling.
 - **Freight Impact:** Freight costs vary significantly across vendors (e.g., ~\$68k for Brown-Forman), which must be factored into net profitability.

4. Technical Implementation & Performance

To ensure this analysis is sustainable for future reporting:

1. **Automation:** A Python script was developed to handle the Extraction, Transformation, and Loading (ETL) process.
2. **Data Cleaning:** Data types were standardized (Volume to float64), missing values were handled (filled with 0), and string padding was removed from categorical names.
3. **Persistence:** The final summary is written back to the database as `vendor_sales_summary`. This prevents the need to re-run expensive joins across 12 million rows every time a report is needed.

5. Strategic Recommendations

- **Vendor Selection:** Use the `SalesToPurchaseRatio` and `ProfitMargin` to renegotiate terms with vendors showing negative or low profitability.
- **Pricing Optimization:** Compare `PurchasePrice` vs. `ActualPrice` vs. `SalesPrice` to identify products that are underpriced relative to their acquisition cost.
- **Inventory Management:** Items with a `StockTurnover` significantly below 1.0 should be reviewed for potential overstocking or markdowns.

Following the Exploratory Data Analysis (EDA) and the subsequent statistical testing, this report synthesizes the findings to provide actionable business insights regarding inventory management, vendor performance, and pricing strategies.

1. Data Cleaning & Statistical Integrity

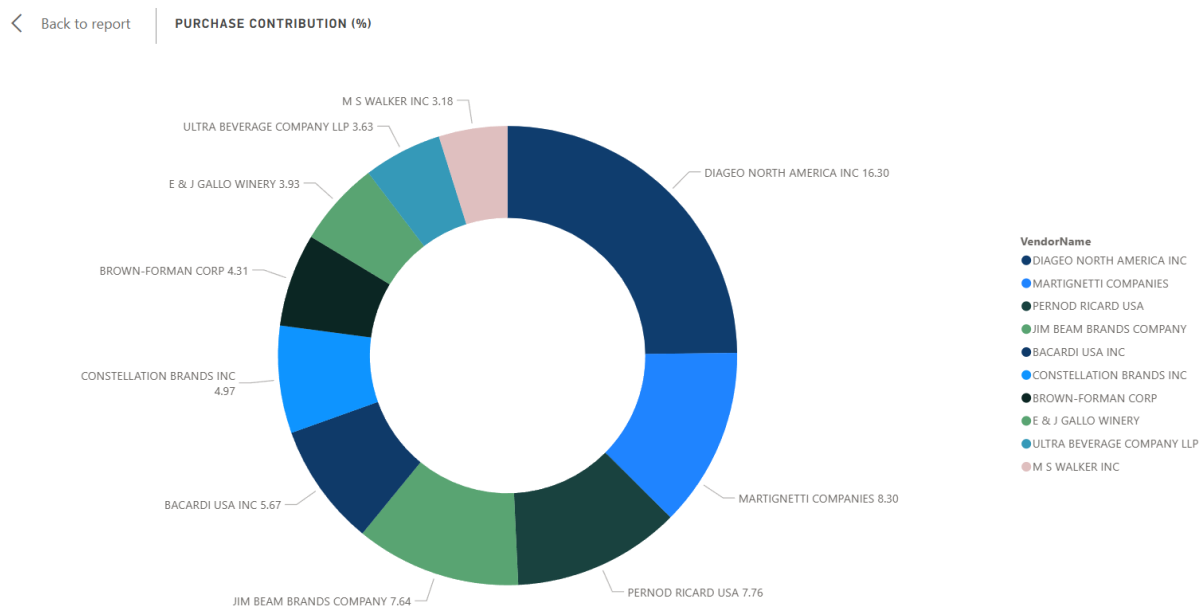
To ensure the reliability of the analysis, the dataset was filtered to remove inconsistencies such as negative gross profits and infinite margins.

- **Final Sample Size:** 8,564 records (brands/vendor combinations).
- **Key Filtering Criteria:** GrossProfit > 0, ProfitMargin > 0, and TotalSalesQuantity > 0.

2. Vendor Performance & Procurement Concentration

The analysis confirms a high level of procurement dependency on a small group of vendors, typical of a **Pareto (80/20) distribution**.

- **Market Dominance:** The top 10 vendors account for **65.69%** of the total purchase contribution.



- **Lead Vendor:** **DIAGEO NORTH AMERICA INC** is the primary contributor, representing **16.3%** of all purchases and generating the highest sales (**\$67.99M**).
- **Capital Allocation:** While these top vendors drive revenue, they also hold the most "locked" capital. **DIAGEO** alone has approximately **\$722K** tied up in unsold inventory.

3. Pricing Strategy & Bulk Purchasing

There is clear evidence that the current bulk purchasing strategy is effective in reducing acquisition costs.

- **Economies of Scale:** * **Large Orders:** ~\$10.78 average unit price.
 - **Small Orders:** ~\$39.07 average unit price.
- **Impact:** A **72% reduction** in unit cost for bulk orders significantly boosts potential gross profit margins, provided that the inventory turnover remains high.

4. The "High Margin, Low Volume" Paradox

Statistical testing (Welch's T-Test and Confidence Intervals) revealed a counter-intuitive relationship between sales volume and profitability.

- **The Findings:** * **Low-Performing Vendors (by sales):** Maintain a significantly higher mean profit margin (**41.57%**).
 - **Top-Performing Vendors (by sales):** Operate on thinner mean margins (**31.18%**).
- **Hypothesis Testing:** With a **P-Value of 0.0000**, we rejected the Null Hypothesis. There is a statistically significant difference in margins between these groups.
- **Business Implication:** High-volume vendors likely utilize "Loss Leader" strategies or competitive pricing to capture market share, while low-volume brands represent "Premium" or niche selections that earn more per unit but move slower.

5. Inventory Efficiency & Risks

- **Stock Turnover:** A strong correlation (**0.999**) between purchase and sales quantities indicates that, generally, the inventory system is highly responsive to demand.
- **Slow-Moving Risks:** Vendors like **ALISA CARR BEVERAGES** (Turnover: **0.61**) and **HIGHLAND WINE MERCHANTS LLC** (Turnover: **0.70**) show signs of overstocking.
- **Capital Exposure:** Total unsold capital across the dataset stands at **\$2.71M**.

6.STATISTICAL SUMMARY:

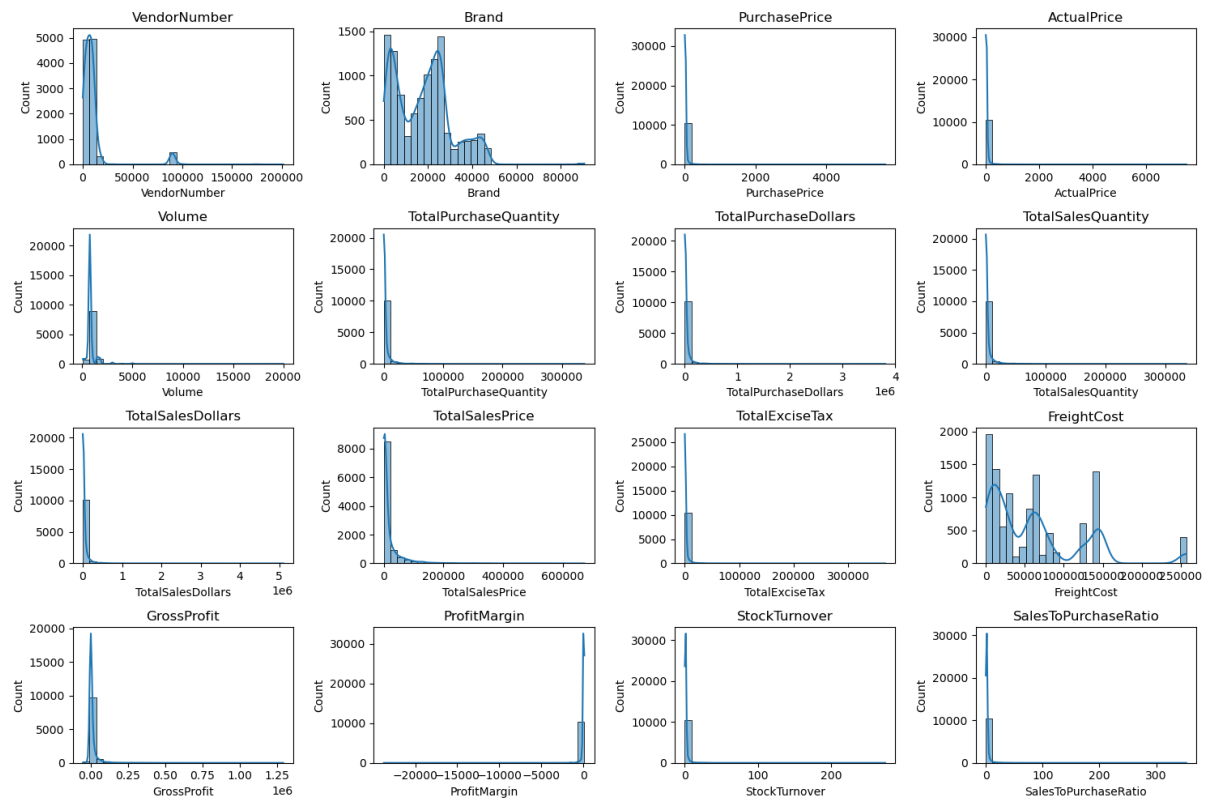
vendor_sales_summary														
	VendorNumber	VendorName	Brand	Description	PurchasePrice	ActualPrice	Volume	TotalPurchaseQuantity	TotalPurchaseDollars	TotalSalesQuantity	TotalSalesDollars	TotalSalesPrice	TotalExciseTax	FreightCost
0	1128	BROWN-FORMAN CORP	1233	Jack Daniels No 7 Black	26.27	36.99	1750	145080	3811251.60	142049.0	5101919.51	672819.31	260999.20	68601.68
1	4425	MARTIGNETTI COMPANIES	3405	Tito's Handmade Vodka	23.19	28.99	1750	164038	3804041.22	160247.0	4819073.49	561512.37	294438.66	144929.24
2	17035	PERNOD RICARD USA	8068	Absolut 80 Proof	18.24	24.99	1750	187407	3418303.68	187140.0	4538120.60	461140.15	343854.07	123780.22
3	3960	DIAGEO NORTH AMERICA INC	4261	Capt Morgan Spiced Rum	16.17	22.99	1750	201682	3261197.94	200412.0	4475972.88	420050.01	368242.80	257032.07
4	3960	DIAGEO NORTH AMERICA INC	3545	Ketel One Vodka	21.89	29.99	1750	138109	3023206.01	135838.0	4223107.62	545778.28	249587.83	257032.07
...
10687	9815	WINE GROUP INC	8537	Concannon Glen Ellen Wh Zin	1.32	4.99	750	2	2.64	5.0	15.95	10.96	0.55	27100.41
10688	8004	SAZERAC CO INC	5683	Dr McGillicuddy's Apple Pie	0.39	0.49	50	6	2.34	134.0	65.66	1.47	7.04	50293.62
10689	3924	HEAVEN HILL DISTILLERIES	9123	Deep Eddy Vodka	0.74	0.99	50	2	1.48	2.0	1.98	0.99	0.10	14069.87
10690	3960	DIAGEO NORTH AMERICA INC	6127	The Club Strawberry Margarita	1.47	1.99	200	1	1.47	72.0	143.28	77.61	15.12	257032.07
10691	7245	PROXIMO SPIRITS INC	3065	Three Olives Grape Vodka	0.71	0.99	50	1	0.71	86.0	85.14	33.66	4.46	38994.78

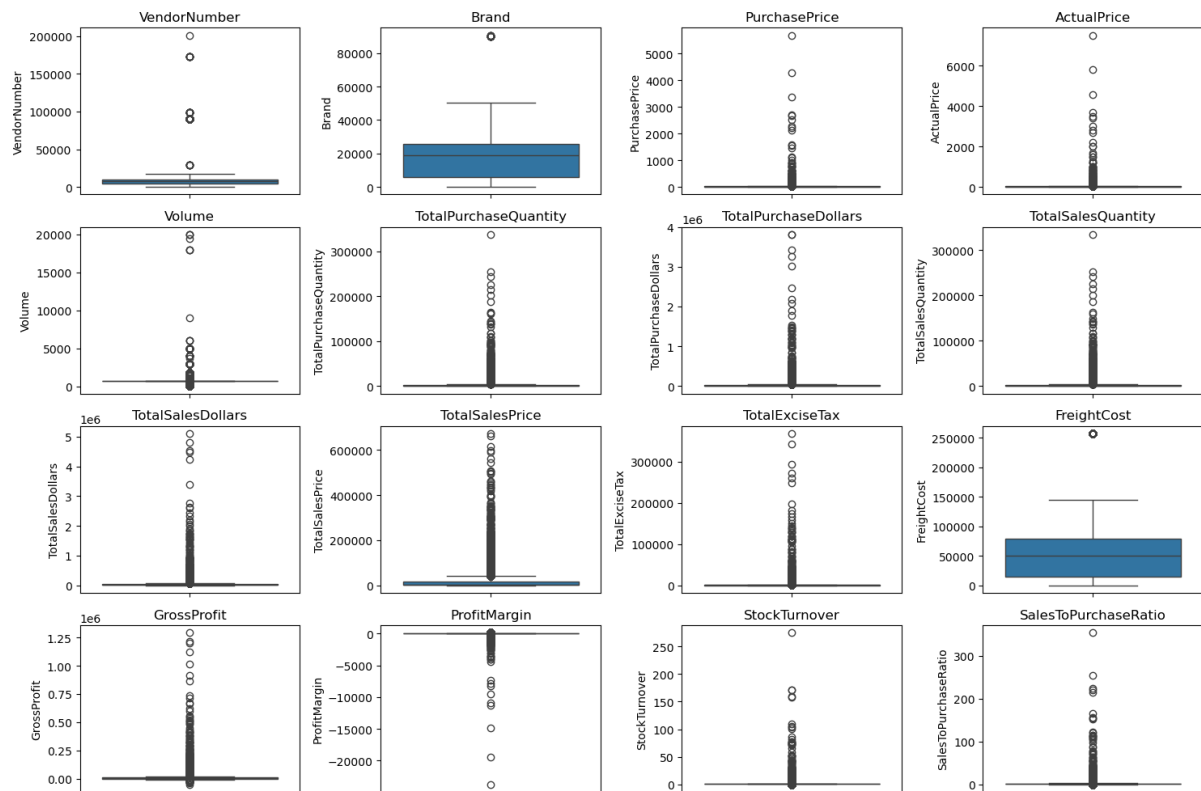
10692 rows × 14 columns

Statistics Insights:

Negative & Zero Values:

- **Gross Profit:** Minimum value is -52,002.78, indicating losses. Some products or transactions may be selling at a loss due to high costs or selling at discounts lower than the purchase price.
- **Profit Margin:** Has a minimum of $-\infty$, which suggests cases where revenue is zero or even lower than costs.
- **Total Sales Quantity & Sales Dollars:** Minimum values are 0, meaning some products were purchased but never sold. These could be slow-moving or obsolete stock.



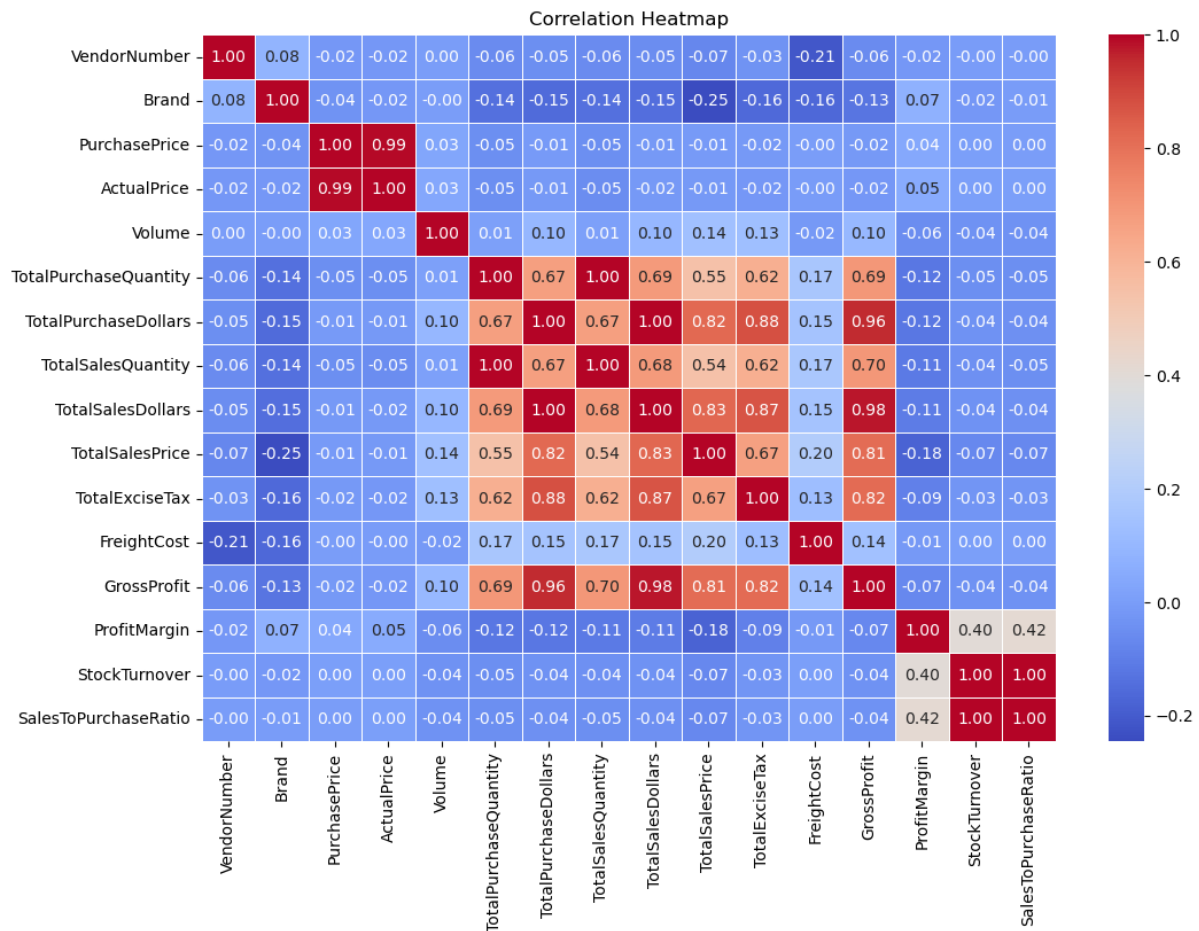


Outliers Indicated by High Standard Deviations:

- **Purchase & Actual Prices:** The max values (5,681.81 & 7,499.99) are significantly higher than the mean (24.39 & 35.64), indicating potential premium products.
- **Freight Cost:** Huge variation, from 0.09 to 257,032.07, suggests logistics inefficiencies or bulk shipments.
- **Stock Turnover:** Ranges from 0 to 274.5, implying some products sell extremely fast while others remain in stock indefinitely. Value more than 1 indicates that sold quantity for that product is higher than purchased quantity due to either sales are being fulfilled from older stock.

Correlation Insights

- **PurchasePrice** has weak correlations with **TotalSalesDollars** (-0.012) and **GrossProfit** (-0.016), suggesting that price variations do not significantly impact sales revenue or profit.
- **Strong correlation** between **total purchase quantity** and **total sales quantity** (0.999), confirming efficient inventory turnover.
- **Negative correlation** between **profit margin** & **total sales price** (-0.179) suggests that as sales price increases, margins decrease, possibly due to competitive pricing pressures.
- **StockTurnover** has weak negative correlations with both **GrossProfit** (-0.038) and **ProfitMargin** (-0.055), indicating that faster turnover does not necessarily result in higher profitability.

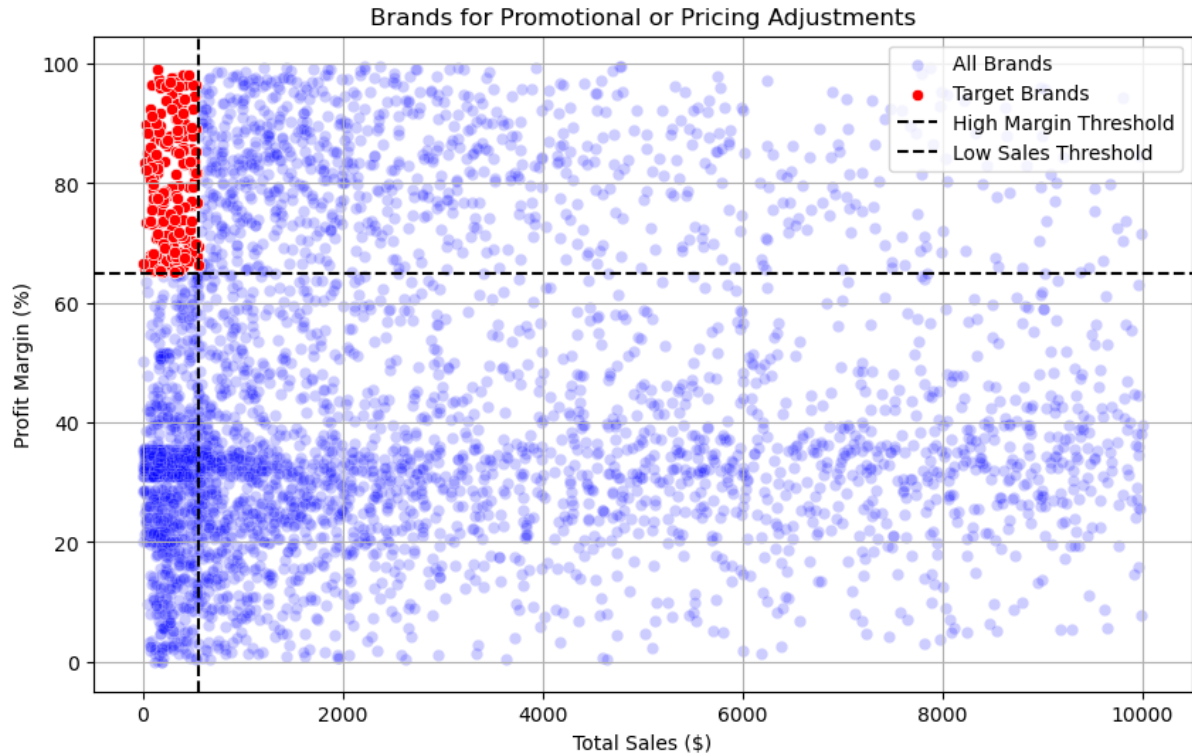


7. Actionable Recommendations

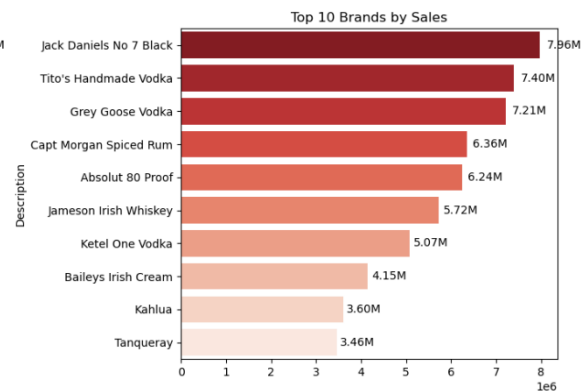
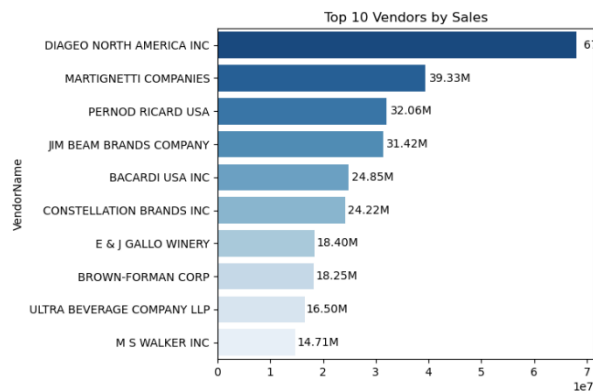
Segment	Strategy	Action Item
Top Vendors	Efficiency	Negotiate better freight rates and focus on reducing the \$2.71M in locked capital.
High Margin Brands	Promotion	Target the 198 brands (e.g., <i>Santa Rita Organic</i> , <i>Crown Royal Apple</i>) that have margins >64% but low sales for marketing campaigns.
Slow Movers	Liquidation	Implement markdowns for vendors with a StockTurnover < 0.80 to free up warehouse space and capital.
Bulk Buyers	Optimization	Continue leveraging the 72% discount on large orders, but only for "A-category" items with high turnover.

Data Analysis

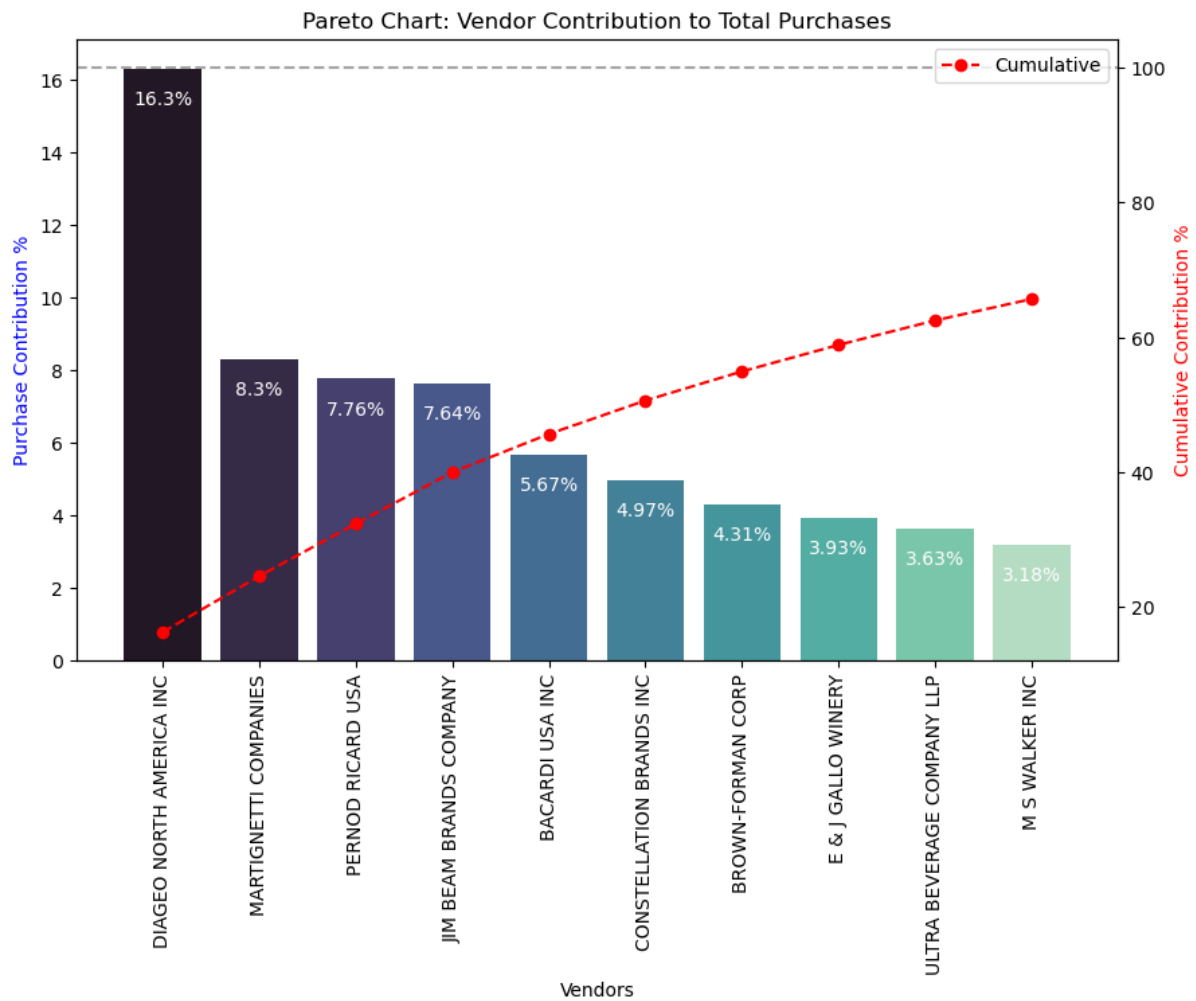
Identify Brands that needs Promotional or Pricing Adjustments which exhibit lower sales performance but higher profit margins.



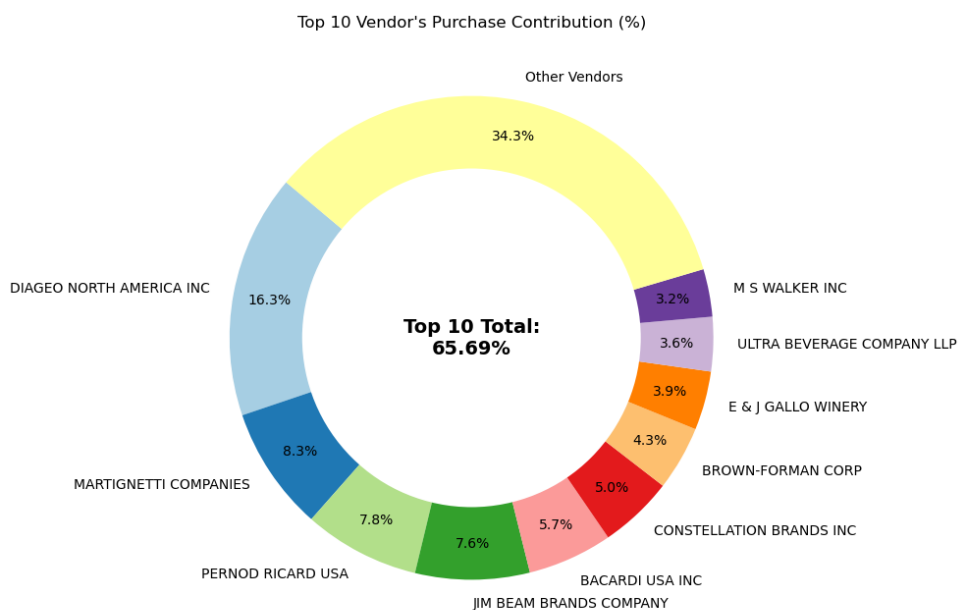
Which vendors and brands demonstrate the highest sales performance?



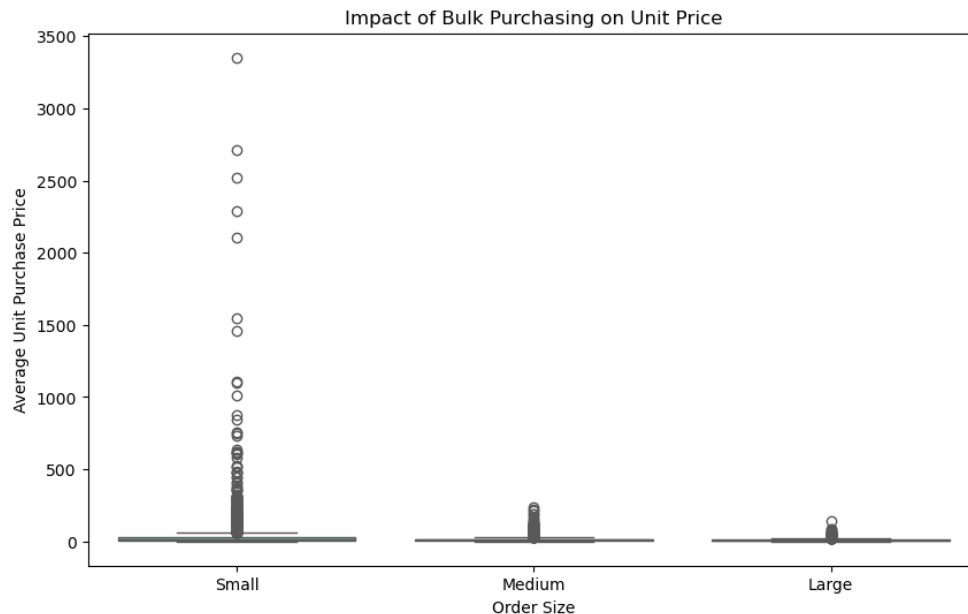
Which vendors contribute the most to total purchase dollars ?



How much of total procurement is dependent on the top vendors?

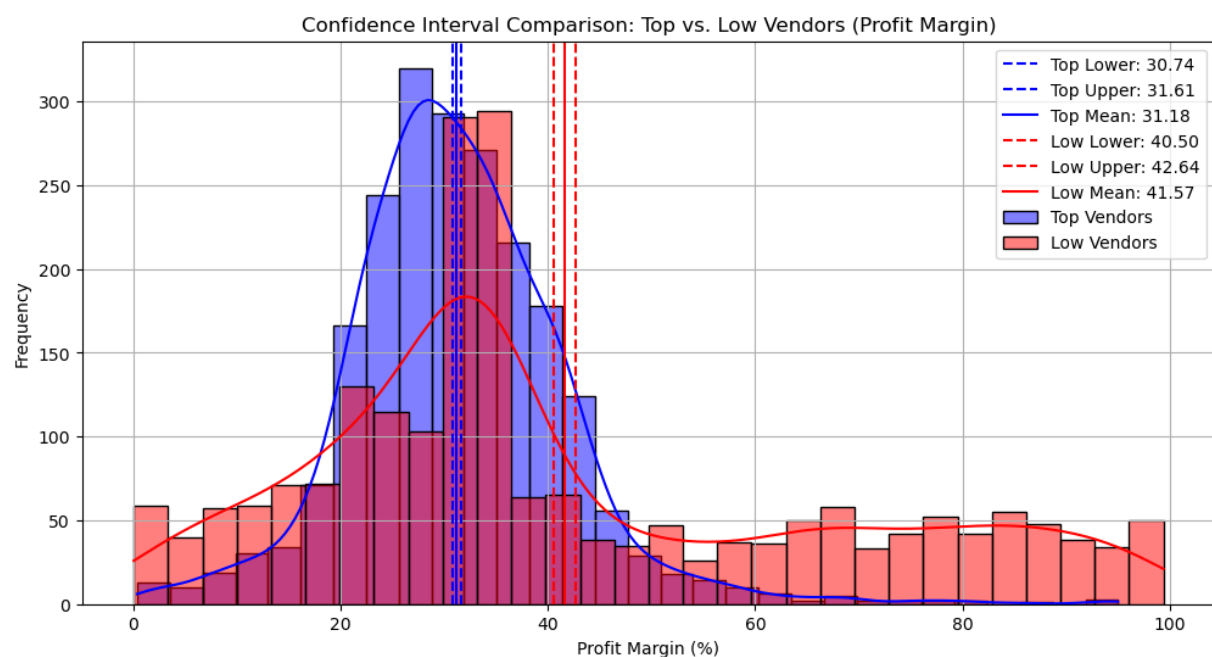


Does purchasing in bulk reduce the unit price, and what is the optimal purchases volume for cost savings?



- Vendors buying in bulk (Large Order Size) get the lowest unit price (\$10.78 per unit), meaning higher margins if they can manage inventory efficiently.
- The price difference between Small and Large orders is substantial (~72% reduction in unit cost)
- This suggests that bulk pricing strategies successfully encourage vendors to purchase in larger volumes, leading to higher overall sales despite lower per-unit revenue.

What is the 95% confidence intervals for profit margins of top-performing and low-performing vendors?



- The **confidence interval** for low-performing vendors (40.48% to 42.62%) is significantly higher than that of top-performing vendors (30.74% to 31.61%).
- This suggests that vendors with **lower sales** tend to maintain **higher profit margins**, potentially due to premium pricing or lower operational costs.
- **For High-Performing Vendors:** If they aim to improve profitability, they could explore selective price adjustments, cost optimization, or bundling strategies.
- **For Low-Performing Vendors:** Despite higher margins, their low sales volume might indicate a need for better marketing, competitive pricing, or improved distribution strategies.

Is there a significant difference in profit margins between top-performing and low-performing vendors?

Hypothesis:

- H_0H_0 (**Null Hypothesis**): There is no significant difference in the mean profit margins of top-performing and low-performing vendors.
- H_1H_1 (**Alternative Hypothesis**): The mean profit margins of top-performing and low-performing vendors are significantly different.

T-Statistic: -17.6695, P-Value: 0.0000 Reject H_0 : There is a significant difference in profit margins between top and low-performing vendors.
--

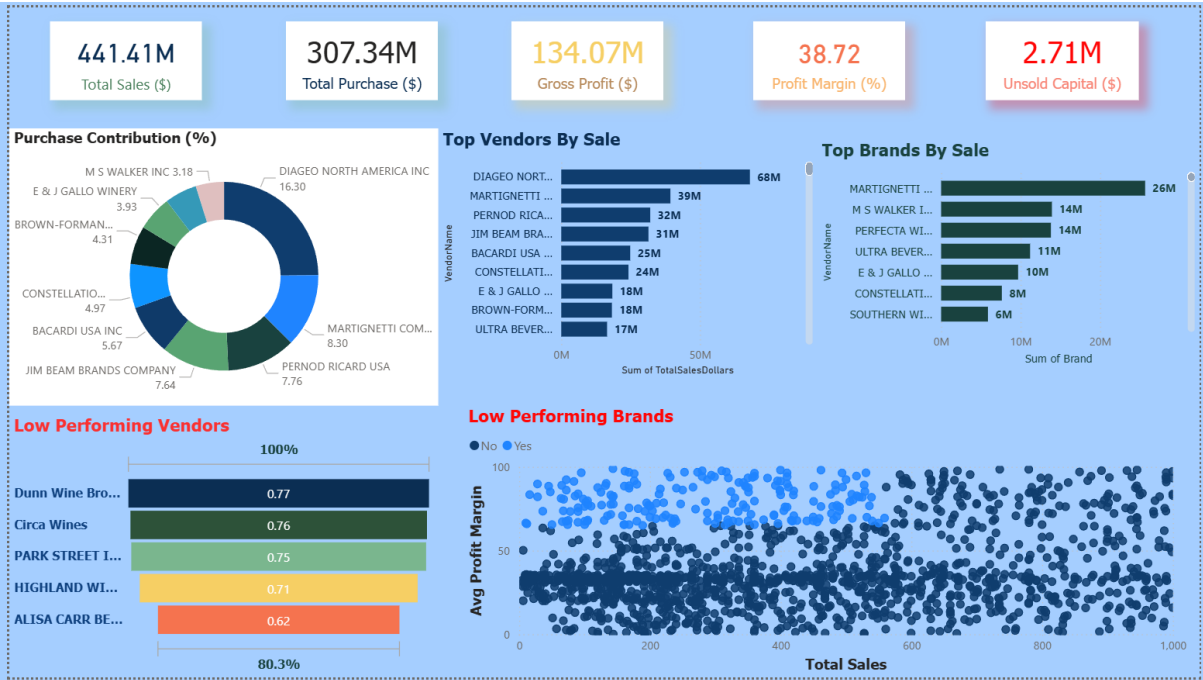
Which vendors have low inventory turnover, indicating excess stock and slow-moving products?

StockTurnover	
VendorName	
ALISA CARR BEVERAGES	0.615385
HIGHLAND WINE MERCHANTS LLC	0.708333
PARK STREET IMPORTS LLC	0.751306
Circa Wines	0.755676
Dunn Wine Brokers	0.766022
CENTEUR IMPORTS LLC	0.773953
SMOKY QUARTZ DISTILLERY LLC	0.783835
TAMWORTH DISTILLING	0.797078
THE IMPORTED GRAPE LLC	0.807569
WALPOLE MTN VIEW WINERY	0.820548

How much capital is locked in unsold inventory per vendor, and which vendddors contribute to most to it?

	VendorName	UnsoldInventoryValue
25	DIAGEO NORTH AMERICA INC	722.21K
46	JIM BEAM BRANDS COMPANY	554.67K
68	PERNOD RICARD USA	470.63K
116	WILLIAM GRANT & SONS INC	401.96K
30	E & J GALLO WINERY	228.28K
79	SAZERAC CO INC	198.44K
11	BROWN-FORMAN CORP	177.73K
20	CONSTELLATION BRANDS INC	133.62K
61	MOET HENNESSY USA INC	126.48K
77	REMY COINTREAU USA INC	118.60K

DASHBOARD (Made in Power Bi)



Conclusion

- This analysis successfully transformed large-scale transactional inventory and sales data into a consolidated, performance-driven Vendor Sales Summary that enables clear, data-backed business decisions. By integrating purchasing, sales, pricing, and freight data, the study provides a holistic view of vendor profitability, inventory efficiency, and pricing effectiveness.
- The findings highlight a **strong dependency on a small group of vendors**, with the top 10 vendors accounting for nearly **66% of total procurement spend**, confirming a classic Pareto concentration. While these vendors drive the majority of revenue, they also account for the **largest share of locked capital**, increasing financial exposure and reinforcing the need for tighter inventory and freight cost controls.
- The pricing and procurement analysis confirms that **bulk purchasing is highly effective**, delivering an average **72% reduction in unit cost**. However, this advantage only translates into profitability when supported by strong inventory turnover. Vendors and brands with slower sales velocity pose a risk of capital lock-in, even when purchase prices are optimized.
- A key strategic insight is the **“high-margin, low-volume” paradox**. Statistical testing confirms a significant difference in profit margins between high- and low-performing vendors. High-volume vendors operate on thinner margins to sustain market share, while low-volume vendors achieve substantially higher margins but suffer from limited demand. This segmentation presents a clear opportunity for **targeted promotions, selective price adjustments, and differentiated vendor strategies** rather than a one-size-fits-all approach.
- Inventory efficiency remains generally strong, as evidenced by the near-perfect correlation between purchase and sales quantities. However, specific vendors with **low stock turnover and high unsold capital** represent immediate optimization opportunities through markdowns, reduced future orders, or supplier renegotiations.
- Overall, this report demonstrates how combining **EDA, statistical validation, and operational metrics** can uncover actionable insights. By acting on these findings—optimizing vendor selection, refining pricing strategies, and reducing excess inventory—the organization can improve profitability, free up working capital, and build a more resilient, data-driven procurement strategy.